Likelihood Ratio Detection and Tracking Tutorial

Presented at Fusion 2002
11 July 2002

Lawrence D. Stone
Metron Inc
11911 Freedom Drive, Suite 800
Reston, VA 20190
Ph: 703 787 8700
Fax: 703 787 3518
email: stone@metsci.com

TABLE OF CONTENTS

1. INTRODUCTION

This tutorial addresses the problem of detection and tracking when there is at most one target present. This problem is most pressing when signal-to-noise ratios are low. This will be the case when one is performing surveillance of a region of the ocean's surface hoping to detect a periscope in the clutter of ocean waves or when scanning the horizon with an infrared sensor trying to detect a cruise missile at the earliest possible moment. Both of these problems have two important features: (1) a target may or may not be present; and (2) if a target is present, it may not produce a strong enough signal to be detected on a single glimpse by the sensor.

Likelihood ratio detection and tracking is based on an extension of Bayesian single target tracking, described in Section 2 below, to the case where there is either one or no target present. The methodology presented here unifies detection and tracking into one seamless process. Likelihood ratio detection and tracking allows both functions to be performed simultaneously and optimally.

2. BAYESIAN SINGLE TARGET TRACKING

In this section we present a Bayesian formulation of single target tracking and a basic recursion for performing single target tracking.

Definition of Bayesian Approach. To appreciate the discussion in this tutorial, the reader must first understand the concept of Bayesian tracking. For a tracking system to be considered Bayesian, it must have the following characteristics.

1. *Prior Distribution*. There must be a prior distribution on the state of the targets. If the targets are moving, the prior distribution has to include a probabilistic description of the motion characteristics of the target. Usually the prior is given in terms of a stochastic process for the motion of the target.

2. *Likelihood Functions*. The information in sensor measurements, observations, or contacts must be characterized by likelihood functions.

3. *Posterior Distribution*. The basic output of a Bayesian tracker is a posterior probability distribution on the (joint) state of the target(s). The posterior at time t is computed by combining the motion updated prior at time t with the likelihood function for the observation(s) received at time t.

These are the basics: prior, likelihood functions, posterior. If these are not present, the tracker is not Bayesian. The recursions given below for performing Bayesian tracking are "recipes" for calculating priors, likelihood functions, and posteriors.

2.1 Bayesian Filtering

Bayesian filtering is based on the mathematical theory of probabilistic filtering described by Jazwinski (1970). Bayesian filtering is the application of Bayesian inference to the problem of tracking a single target. In this section, we consider the situation where the target motion is modeled in continuous time, but the observations are received at discrete, possibly random, times. This is called continuous-discrete filtering by Jazwinski.

2.2 Problem Definition

The single target tracking problem assumes that there is one target present in the state space; as a result, the problem becomes one of estimating the state of that target.

2.2.1 Target State Space

Let S be the state space of the target. Typically, the target state will be a vector of components. Usually some of these components are kinematic and include position, velocity, and possibly acceleration. Note that there may be constraints on the components, such as a maximum speed for the velocity component. There can be additional components that may be related to the identity or other features of the target. For example, if one of the components specifies target type, then that may also specify information such as radiated noise levels at various frequencies and motion characteristics (e.g., maximum speeds). In order to use the recursion presented in this section, there are additional requirements on the target state space. The state space must be rich enough that (1) the target's motion is Markovian in the chosen state space and (2) the sensor likelihood functions depend only on the state of the target at the time of the observation.

The sensor likelihood functions depend on the characteristics of the sensor, such as its position and measurement error distribution which are assumed to be known. If they are not known, they need to be determined by experimental or theoretical means.

2.2.2 Prior Information

Let $X(t)$ be the (unknown) target state at time *t*. We start the problem at time 0 and are interested in estimating $X(t)$ for $t \geq 0$. The prior information about the target is represented by a stochastic process $\{X(t); t \geq 0\}$. Sample paths of this process correspond to possible target paths through the state space, *S*. The state space S has a measure associated with it. If S is discrete, this measure is a discrete measure. If S is continuous (e.g., if S is equal to the plane), then this measure is represented by a density. The measure on S can be a mixture or product of discrete and continuous measures. Integration with respect to this measure will be indicated by *ds*. If the measure is discrete, then integration becomes summation.

Markov Target Motion Assumption. In order to calculate the posterior distributions in a recursive manner we shall assume that the target's motion is Markovian in the state space *S*. With this in mind we define the transition function

$$q_k(s_k \mid s_{k-1}) = \mathbf{Pr}\{X(t_k) = s_k \,\big|\, X(t_{k-1}) = s_{k-1}\} \text{ for } k \geq 1,$$

and let q_0 be the probability (density) function for *X*(0). By the Markov assumption

$$\mathbf{Pr}\left\{X(t_1) = s_1, \ldots, X(t_K) = s_K\right\} = \int_S \prod_{k=1}^{K} q_k(s_k \mid s_{k-1})\, q_0(s_0)\, ds_0 \ . \tag{1}$$

2.2.3 Sensors

There is a set of sensors that report observations at an ordered, discrete sequence of (possibly random) times. These sensors may be of different types and report different information. The set can include radar, sonar, infra-red, visual, and other types of sensors. The sensors may report only when they have a contact or on a regular basis. Observations from sensor *j* take values in the measurement space H_j. Each sensor may have a different measurement space. The probability distribution of each sensor's response conditioned on the value of the target state *s* is assumed to be known. This relationship is captured in the likelihood function for that sensor.

2.2.4 Likelihood Function

Suppose that we receive an observation (measurement) *y* from sensor *j* at time *t*. We assume that we have a model of the sensor performance and its error characteristics that allows us to compute the probability (density) of the sensor observations as a function of target state. Specifically we assume that the senor observation *Y* is random variable whose distribution conditioned on target state *s* is known. Thus we can compute

$$L(y \mid s) \equiv \mathbf{Pr}\{Y = y \mid X(t) = s\} \text{ for } s \in S\ .$$

$L(y \mid \cdot)$ is defined to be the *likelihood function* for measurement *y*. Notice that *L* is a function on the target state space *S* regardless of the measurement space of the sensor. *Thus likelihood functions become the common currency of information in a Bayesian Tracker*. They replace and generalize the notion of contacts. Likelihood functions can represent sensor information such as detections, no detections, Gaussian contacts, bearing observations, measured signal-to-noise ratios, and observed frequencies of a signal. Likelihood functions can represent and incorporate information in situations where the notion of a contact is not meaningful. Subjective information also can be incorporated by using likelihood functions. Examples of likelihood functions are provided in Section 2.4.

Conditional Independence. Suppose that by time *t* we have obtained observations at the set of times $0 \leq t_1 \leq \ldots \leq t_K \leq t$. To allow for the possibility that we may receive more than one sensor observation at a given time, we let Y_k be the set of sensor observations received at time t_k. Let y_k denote a value of the random variable Y_k. We assume that we can compute the likelihood function

$$L_k(y_k \mid s) = \mathbf{Pr}\{Y_k = y_k \mid X(t_k) = s\} \text{ for } s \in S \text{ and } k = 1, \ldots, K . \tag{2}$$

Let $\mathbf{Y}(t) = (Y_1, Y_2, \ldots, Y_K)$ and $\mathbf{y} = (y_1, \ldots, y_K)$. Define

$$L(\mathbf{y} \mid s_1, \ldots, s_K) = \mathbf{Pr}\left\{\mathbf{Y}(t) = \mathbf{y} \mid X(t_1) = s_1, \ldots, X(t_K) = s_K\right\}.$$

We further assume that

$$\mathbf{Pr}\left\{\mathbf{Y}(t) = \mathbf{y} \mid X(u) = s(u), 0 \le u \le t\right\} = L(\mathbf{y} \mid s_1, \ldots, s_K) = \prod_{k=1}^{K} L_k(y_k \mid s_k) . \tag{3}$$

Equation (3) means that the likelihood of the data $\mathbf{Y}(t)$ received through time t depends only on the target states at the times $\{t_1, \ldots, t_K\}$ and not on the whole target path. Furthermore, the likelihood function of the observation at time t_k is independent of the observations at all other times given $X(t_k) = s_k$. This is the conditional independence assumption made for Bayesian recursive tracking.

2.2.5 Posterior on Target State

Let

$$p(t_K, s_K) = \mathbf{Pr}\left\{X(t_K) = s_K \mid \mathbf{Y}(t_K) = \mathbf{y}\right\}.$$

Note that the dependence of p on $\mathbf{y}$ has been suppressed. The function $p(t_K, \cdot)$ is the posterior distribution on $X(t_K)$ given $\mathbf{Y}(t_K) = \mathbf{y}$. In mathematical terms, the problem is to compute this posterior distribution. Recall that from the point of view of Bayesian inference, the posterior distribution on target state represents our knowledge of the target state. All estimates of target state derive from this posterior.

2.3 Computing the Posterior

By the Markov and conditional independence assumptions, we may compute the posterior using the following recursion.

Basic Recursion for Single Target Tracking

Initialize Distribution: $p(t_0, s_0) = q_0(s_0)$ for $s_0 \in S$ (4)

For $k \geq 1$ and $s_k \in S$,

Perform Motion Update: $p^-(t_k, s_k) = \int q_k(s_k \mid s_{k-1})\, p(t_{k-1}, s_{k-1}) ds_{k-1}$ (5)

Compute Likelihood Function L_k from the observation $Y_k = y_k$.

Perform Information Update: $p(t_k, s_k) = \frac{1}{C} L_k(y_k \mid s_k)\, p^-(t_k, s_k)$ (6)

The motion update in (5) accounts for the transition of the target state from time t_{k-1} to t_k . Transitions can represent not only the physical motion of the target, but also changes in other state variables. The information update in (6) is accomplished by pointwise multiplication of $p^-(t_k, s_k)$ by the likelihood function $L_k(y_k \mid s_k)$. If there has been no observation at time t_k , then there is no information update, only a motion update.

Except in special circumstances, this recursion must be computed numerically. Today's high-powered scientific workstations can compute and display tracking solutions for complex nonlinear trackers. One way to do this is to discretize the state space and use a Markov chain model for target motion so that (5) is computed through the use of discrete transition probabilities. The likelihood functions are also computed on the discrete state space. A numerical implementation of a discrete Bayesian tracker is described in section 3.3 of Stone *et al*. (1999). A more common numerical implementation is to use a particle filter as described in Doucet, de Freitas, and Gordon (2001).

2.4 Examples of Likelihood Functions

In the classical view of tracking, contacts are obtained from sensors that provide estimates of (some components of) the target state at a given time with a specified measurement error. In the Kalman filter formulation, a measurement (contact) Y_k at time t_k satisfies the measurement equation

$$Y_k = \mathbf{M}_k X(t_k) + \varepsilon_k \tag{7}$$

where

Y_k is an *r*-dimensional real column vector

$X(t_k)$ is an *l*-dimensional real column vector

$\mathbf{M}_k$ is an $r \times l$ matrix

$\varepsilon_k \sim \mathcal{N}(0, \mathbf{\Sigma}_k)$.

Note that $\sim \mathcal{N}(\mu, \mathbf{\Sigma})$ means "has a Normal (Gaussian) distribution with mean μ and covariance $\mathbf{\Sigma}$." In this case, the measurement is a linear function of the target state and the measurement error is Gaussian.

2.4.1 Gaussian Contact Likelihood Function

We can express a Gaussian contact in terms of a likelihood function as follows. Let $L_G(y \mid x) = \Pr\{Y_k = y \mid X(t_k) = x\}$. Then

$$L_G\left(y \mid x\right) = \left(2\pi\right)^{-r/2} \left|\det \mathbf{\Sigma}_k\right|^{-1/2} \exp\left(-\frac{1}{2}\left(y - \mathbf{M}_k x\right)^T \Sigma^{-1}\left(y - \mathbf{M}_k x\right)\right). \tag{8}$$

Note that the measurement *y* is data that is known and fixed. The target state *x* is unknown and varies, so that the likelihood function is a function of the target state variable *x*. Equation (8) looks the same as a standard elliptical contact, or estimate of target state, expressed in the form of multivariate normal distribution, commonly used in Kalman filters. There is a difference, but it is obscured by the symmetrical positions of *y* and $\mathbf{M}_k x$ in the Gaussian density in (8). A likelihood function does not represent an estimate of the target state. It looks at the situation in reverse. For each value of target state *x*, it calculates the probability (density) of obtaining the measurement *y* given that the target is in state *x*. In most cases, likelihood functions are not probability (density) functions on the target state space. They need not integrate to one over the target state space. In fact, the likelihood function in (8) is a probability density on the target state space only when Y_k is an *l*-dimensional and $\mathbf{M}_k$ is an $l \times l$ matrix.

2.4.2 Line of Bearing Plus Detection Likelihood Functions

Suppose that there is a sensor located in the plane at (70,0) and that it has produced a detection. For this sensor the probability of detection is a function, $P_d(r)$, of the range *r* from the sensor. Take the case of an underwater sensor such as an array of acoustic hydrophones and a situation where the propagation conditions produce convergence zones of high detection performance that alternate with ranges of poor detection performance. The observation (measurement) in this case is $Y = 1$ for detection and 0 for no detection. The likelihood function for detection is

$$L_d(1 \mid x) = P_d\left(r(x)\right), \text{ where } r(x) \text{ is the range from the state } x \text{ to the sensor.}$$

Figure 1 shows the likelihood function for this observation.

Suppose that, in addition to the detection, there is a bearing measurement of 135 degrees (measured counter-clockwise from the x_1 axis) with a Gaussian measurement error having mean 0 and standard deviation 15 degrees. Figure 2 shows the likelihood function for this observation. Notice that, although the measurement error is Gaussian in bearing, it does not produce a Gaussian likelihood function on the target state space. Furthermore, this likelihood function would integrate to infinity over the whole state space. The information from these two likelihood functions is combined by pointwise multiplication. Figure 3 shows the likelihood function that results from this combination.

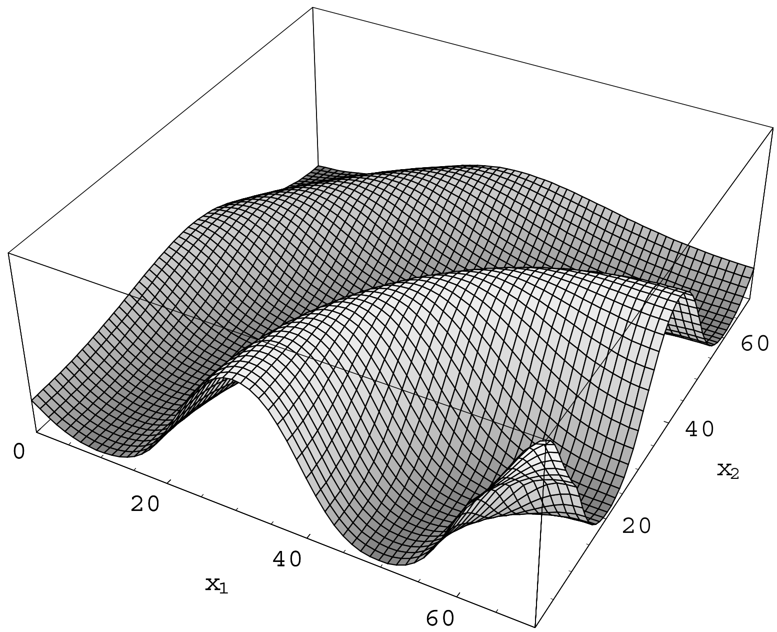

Figure 1. Detection Likelihood Function for a Sensor at (70,0)

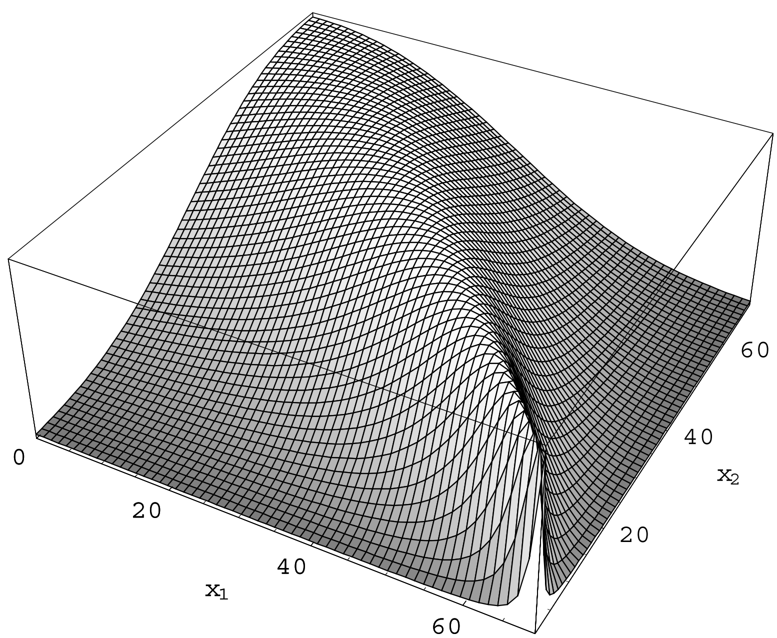

Figure 2. Bearing Likelihood Function for a Sensor at (70,0)

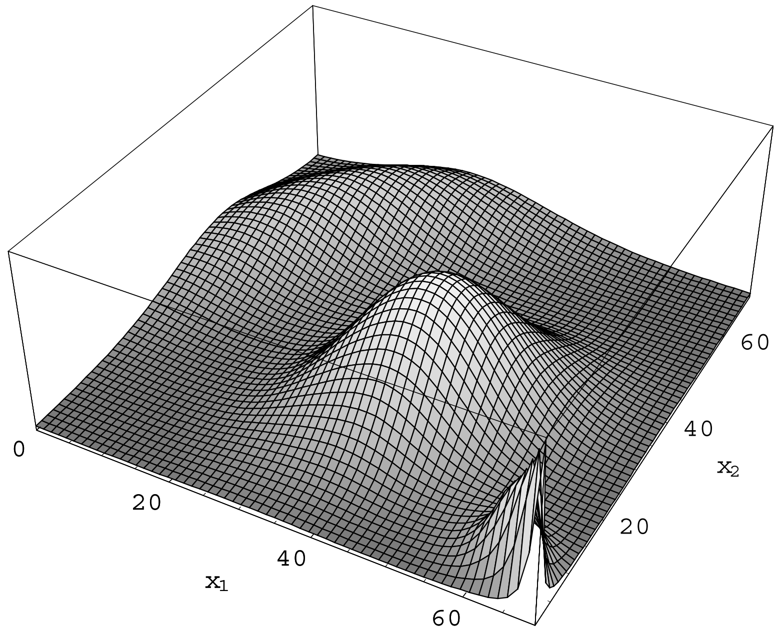

Figure 3. Combined Bearing and Detection Likelihood Function

2.4.3 Combining Information Using Likelihood Functions

Although the example of combining likelihood functions presented in section 2.4.2 is simple, it illustrates the power of using likelihood functions to represent and combine information. A likelihood function converts the information in a measurement to a function on the target state space. Since all information is represented on the same state space, it can easily and correctly be combined, regardless of how disparate the sources of the information. The only limitation is the ability to compute the likelihood function corresponding to the measurement or the information to be incorporated. As an example, subjective information can often be put into the form of a likelihood function and incorporated into a tracker if desired.

3. LIKELIHOOD RATIO DETECTION AND TRACKING

This section describes the problem of detection and tracking when there is at most one target present.

3.1 Basic Definitions and Relations

We use the same assumptions as in Section 2, with the following difference. We do not assume a target is present. Instead the target state space S is augmented with a null state ϕ to make $S^+ = S \cup \phi$. The null state represents no target present. There is a probability (density) function, p, defined on S^+ such that

$$p(\phi) + \int_{s \in S} p(s)ds = 1.$$

Both the state of the target $X(t) \in S^+$ and the information accumulated for estimating the state probability densities evolve with time *t*. The process of target detection and tracking consists of computing the posterior version of the function p as new observations are available and propagating it to reflect the temporal evolution implied by target dynamics. Target dynamics include the probability of target motion into and out of *S* as well as the probabilities of target state changes.

Following the notation used in Section 2 for single-target Bayesian filtering, we let

$$p(t,s) = \mathbf{Pr}\left\{X(t) = s | \mathbf{Y}(t) = \left(Y(t_1),\ldots,Y(t_K)\right)\right\} \text{ for } s \in S^+$$

so that $p(t,\cdot)$ is the posterior distribution on $X(t)$ given all observations received through time *t*. Recall that

$$p^-(t_k, s_k) = \int_{S^+} q(s_k \mid s_{k-1})\, p(t_{k-1}, s_{k-1})\, ds_{k-1} \text{ for } s_k \in S^+$$

is the posterior from time t_{k-1} updated for target motion to time t_k, the time of the *k*th observation. Recall also the definition of the *likelihood* function L_k. Specifically, for the observation $Y_k = y_k$

$$L_k(y_k \mid s) = \mathbf{Pr}\left\{Y_k = y_k \mid X(t_k) = s\right\} \tag{9}$$

where for each $s \in S^+$, $L_k(\cdot \mid s)$ is a probability (density) function on the measurement space H_k.

According to Bayes' Rule,

$$\begin{aligned} p(t_k, s) &= \frac{p^-(t_k, s) L_k(y_k \mid s)}{C(k)} \quad \text{for } s \in S \\ p(t_k, \phi) &= \frac{p^-(t_k, \phi) L_k(y_k \mid \phi)}{C(k)}. \end{aligned} \tag{10}$$

In these equations, the denominator is the likelihood of obtaining the measurement $Y_k = y_k$, that is,

$$C(k) = p^-(t_k, \phi) L_k(y_k \mid \phi) + \int_{s \in S} p^-(t_k, s) L_k(y_k \mid s)\, ds\,.$$

3.1.1 Target Likelihood Ratio

The ratio of the state probability (density) to the null state probability $p(\phi)$ is defined to be the *target likelihood ratio (density)*, $\Lambda(s)$; that is,

$$\Lambda(s) = \frac{p(s)}{p(\phi)} \text{ for } s \in S . \tag{11}$$

The notation for Λ is consistent with that already adopted for the probability densities. Thus, the prior and posterior forms become

$$\Lambda^-(t,s) = \frac{p^-(t,s)}{p^-(t,\phi)} \text{ and } \Lambda(t,s) = \frac{p(t,s)}{p(t,\phi)} \text{ for } s \in S \text{ and } t \geq 0 . \tag{12}$$

The target likelihood ratio density has the same dimensions as the state probability density. Furthermore, from the target likelihood ratio density one may easily recover the state probability density as well as the probability of the null state. Since

$$\int_S \Lambda(t,s)\, ds = \frac{1 - p(t,\phi)}{p(t,\phi)},$$

it follows that

$$\begin{aligned} p(t,s) &= \frac{\Lambda(t,s)}{1 + \int_S \Lambda(t,s')ds'} \text{ for } s \in S \\ p(t,\phi) &= \frac{1}{1 + \int_S \Lambda(t,s')ds'}. \end{aligned} \tag{13}$$

3.1.2 Measurement Likelihood Ratio

The measurement likelihood ratio $\mathcal{L}_k$ for the observation Y_k is defined as

$$\mathcal{L}_k(y \mid s) = \frac{L_k(y \mid s)}{L_k(y \mid \phi)} \text{ for } y \in H_k,\ s \in S . \tag{14}$$

$\mathcal{L}_k(y \mid s)$ is the ratio of the likelihood of receiving the observation $Y_k = y_k$ (given the target is in state s) to the likelihood of receiving $Y_k = y_k$ given no target present. As discussed by Van Trees (1967) the measurement likelihood ratio has long been recognized as part of the prescription for optimal receiver design.

Measurement likelihood ratio functions are chosen for each sensor to reflect its salient properties such as noise characterization and target effects. These functions contain all the sensor information that is required for making optimal Bayesian inferences from sensor measurements.

3.2 Likelihood Ratio Recursion

Under the assumptions given in Section 2 for single-target tracking, the following recursion for calculating the target likelihood ratio holds.

Likelihood Ratio Recursion

Initialize $$p(t_0, s) = q_0(s) \text{ for } s \in S^+ \tag{15}$$

For $k \geq 1$ and $s \in S^+$,

Perform Motion Update $$p^-(t_k, s) = \int_{S^+} q_k(s \mid s_{k-1})\, p(t_{k-1}, s_{k-1}) ds_{k-1} \tag{16}$$

Calculate Likelihood Function $$L_k(y_k \mid s) = \mathbf{Pr}\left\{Y_k = y_k \mid X(t_k) = s\right\} \tag{17}$$

Perform Information Update $$p(t_k, s) = \frac{1}{C} L_k(y_k \mid s)\, p^-(t_k, s) \tag{18}$$

For $k \geq 1$,

Calculate Target Likelihood Ratio $$\Lambda(t_k, s) = \frac{p(t_k, s)}{p(t_k, \phi)} \text{ for } s \in S\,. \tag{19}$$

The constant, C, in (18) is a normalizing factor that makes $p(t_k, \cdot)$ a probability (density) function.

Simplified Recursion. The recursion given in Equations (15) - (19) requires the computation of the full probability function $p(t_k, \cdot)$ using the basic recursion for single target tracking discussed in Section 2. A simplified version of the likelihood ratio recursion has probability mass flowing from the state ϕ to S and from S to ϕ in such a fashion that

$$\begin{aligned} p^-(t_k, \phi) &= q_k(\phi \mid \phi)\, p(t_{k-1}, \phi) + \int_S q_k(\phi \mid s) p(t_{k-1}, s)\, ds \\ &= p(t_{k-1}, \phi). \end{aligned} \tag{20}$$

Since

$$p^-(t_k, s_k) = q_k(s_k \mid \phi)\, p(t_{k-1}, \phi) + \int_S q_k(s_k \mid s) p(t_{k-1}, s)\, ds \text{ for } s_k \in S\,,$$

we have

$$\begin{aligned} \Lambda^-(t_k, s_k) &= \frac{q_k(s_k \mid \phi)\, p(t_{k-1}, \phi) + \int_S q_k(s_k \mid s) p(t_{k-1}, s)\, ds}{p^-(t_k, \phi)} \\ &= \frac{q_k(s_k \mid \phi) + \int_S q_k(s_k \mid s)\, \Lambda(t_{k-1}, s)\, ds}{p^-(t_k, \phi) / p(t_{k-1}, \phi)}. \end{aligned}$$

From (20) it follows that

$$\Lambda^{-}(t_k, s_k) = q_k(s_k \mid \phi) + \int_S q_k(s_k \mid s)\,\Lambda(t_{k-1}, s)\,ds \quad \text{for } s_k \in S\ . \tag{21}$$

Assuming (20) holds, we can write a simplified version of the basic likelihood ratio recursion

Simplified Likelihood Ratio Recursion

Initialize Likelihood Ratio $$\Lambda(t_0, s) = \frac{p(t_0, s)}{p(t_0, \phi)} \quad \text{for } s \in S \tag{22}$$

For $k \geq 1$ and $s \in S$,

Perform Motion Update $$\Lambda^{-}(t_k, s) = q_k(s \mid \phi) + \int_S q_k(s \mid s_{k-1})\,\Lambda(t_{k-1}, s_{k-1})ds_{k-1} \tag{23}$$

Calculate Measurement Likelihood Ratio $$\mathcal{L}_k(y \mid s) = \frac{L_k(y \mid s)}{L_k(y \mid \phi)} \tag{24}$$

Perform Information Update $$\Lambda(t_k, s) = \mathcal{L}_k(y_k \mid s)\,\Lambda^{-}(t_k, s) \tag{25}$$

The simplified recursion is a reasonable approximation to problems involving surveillance of a region that may or may not contain a target. Targets may enter and leave this region, but only one target is in the region at a time.

As a special case, consider the situation where no mass moves from state ϕ to S or from S to ϕ under the motion assumptions. In this case $q_k(s \mid \phi) = 0$ for all $s \in S$ and $p^{-}(t_k, \phi) = p(t_{k-1}, \phi)$ so that (23) becomes

$$\Lambda^{-}(t_k, s) = \int_S q_k(s \mid s_{k-1})\,\Lambda(t_{k-1}, s_{k-1})ds_{k-1}\ . \tag{26}$$

3.3 Log-Likelihood Ratios

Frequently it is more convenient to write (25) in terms of natural logarithms. Doing so results in quantities that require less numerical range for their representation. Another advantage is that, frequently, the logarithm of the measurement likelihood ratio is a simpler function of the observations than is the actual measurement likelihood ratio itself. For example, when the measurement consists of an array of numbers, the measurement log-likelihood ratio often becomes a linear combination of those data, whereas the measurement likelihood ratio involves a product of powers of the data. In terms of logarithms, (25) becomes

$$\ln \Lambda(t_k, s) = \ln \Lambda^{-}(t_k, s) + \ln \mathcal{L}_k(y_k \mid s) \quad \text{for } s \in S\ . \tag{27}$$

The following example is provided to impart an understanding of the practical differences between a formulation in terms of probabilities and a formulation in terms of the logarithm of the likelihood ratios. Suppose there are *I* discrete target states, corresponding to physical locations so that the target state $X \in \{s_1, s_2, ..., s_I\}$ when the target is present. The observation is a vector, $\mathbf{Y}$, that is formed from measurements corresponding to these spatial locations, so that $\mathbf{Y} = (Y(s_1), ..., Y(s_I))$. In the absence of a target in state, s_i, the observation $Y(s_i)$ has a distribution with density function $\eta(\cdot, 0, 1)$, where $\eta(\cdot, \mu, \sigma^2)$ is the density function for a Gaussian distribution with mean μ and variance σ^2. The observations are independent of one another regardless of whether a target is present. When a target is present in the *i*th state, the mean for $Y(s_i)$ is shifted from 0 to a value *r*. In order to perform a Bayesian update, the likelihood function for the observation $\mathbf{Y} = \mathbf{y} = (y(s_1), \ldots, y(s_I))$ is computed as follows:

$$\begin{aligned} L(\mathbf{y} \mid s_i) &= \eta\left(y(s_i), r, 1\right) \prod_{j \neq i} \eta\left(y(s_j), 0, 1\right) \\ &= \exp(r y(s_i) - \tfrac{1}{2} r^2) \prod_{j=1}^{I} \eta\left(y(s_j), 0, 1\right). \end{aligned}$$

Contrast this with the form of the measurement log-likelihood ratio for the same problem. For state *i*, we have

$$\ln \mathcal{L}(\mathbf{y} \mid s_i) = r y(s_i) - \tfrac{1}{2} r^2 .$$

Fix s_i and consider $\ln \mathcal{L}(\mathbf{Y} \mid s_i)$ as a random variable. That is, consider $\ln \mathcal{L}(\mathbf{Y} \mid s_i)$ before making the observation. It has a Gaussian distribution with

$$\begin{aligned} \mathbf{E}\left[\ln \mathcal{L}(\mathbf{Y} \mid s_i) \mid X = s_i\right] &= +\tfrac{1}{2} r^2 \\ \mathbf{E}\left[\ln \mathcal{L}(\mathbf{Y} \mid s_i) \mid X = \phi\right] &= -\tfrac{1}{2} r^2 \\ \mathbf{Var}\left[\ln \mathcal{L}(\mathbf{Y} \mid s_i)\right] &= r^2. \end{aligned}$$

This reveals a characteristic result. *Whereas the likelihood function for any given state requires examination and processing of all the data, the measurement log-likelihood ratio for a given state commonly depends on only a small fraction of the data — frequently only a single datum.* Typically, this will be the case when the observation **Y** is a vector of independent observations.

Because of the importance of the logarithm of the target likelihood ratio density in this methodology, playing as large a role as the target likelihood ratio density itself, we shall assign a special symbol to it. We define $\lambda = \ln \Lambda$ in all of its variations, employing λ^- to refer to the prior and leaving the unadorned symbol λ to refer to the posterior value. Equation (27) assumes the following form in this notation:

$$\lambda(t_k, s) = \lambda^-(t_k, s) + \ln \mathcal{L}_k(y_k \mid s) \quad \text{for } s \in S \tag{28}$$

Because the target likelihood ratio density is not a scalar but has dimensions of the inverse of the state space volume, the quantity λ depends upon the units chosen to describe the state space of the target.

3.4 Declaring a Target Present

The likelihood ratio methodology allows the Bayesian posterior probability density to be computed, including the discrete probability that no target resides in *S* at a given time. It extracts all possible inferential content from the knowledge of the target dynamics, the apriori probability structure, and the evidence of the sensors. This probability information may be used in a number of ways to decide whether a target is present. The following offers a number of traditional methods for making this decision, all based on the integrated target likelihood ratio. Define

$$p(t,1) = \int_S p(t,s)ds = \mathbf{Pr}\{\text{target present in } S \text{ at time } t\}.$$

Then

$$\bar{\Lambda}(t) = p(t,1) / p(t,\phi)$$

is defined to be the *integrated target likelihood ratio at time t*. It is the ratio of the probability of the target being present in *S* to the probability of the target not being present in *S* at time *t*.

Minimizing Bayes' risk. To calculate Bayes' risk, costs must be assigned to the possible outcomes related to each decision (e.g., declaring a target present or not). Define the following costs:

$C(1|1)$ if we declare the target present and it is present
$C(1|\phi)$ if we declare the target present and it is not present
$C(\phi|1)$ if we declare the target not present and it is present
$C(\phi|\phi)$ if we declare the target not present and it is not present.

We assume that it is always better to declare the correct state; that is,

$$C(1|1) < C(\phi|1) \text{ and } C(\phi|\phi) < C(1|\phi).$$

The *Bayes' risk* of a decision is defined as the expected cost of making that decision. Specifically the Bayes' risk is

$p(t,1)C(1|1) + p(t,\phi)C(1|\phi)$ for declaring a target present

$p(t,1)C(\phi|1) + p(t,\phi)C(\phi|\phi)$ for declaring a target not present.

One procedure for making a decision is to take that action which minimizes the Bayes' risk. Applying this criterion produces the following decision rule. Define the threshold

$$\Lambda_{\mathrm{T}} = \frac{C(1|\phi) - C(\phi|\phi)}{C(\phi|1) - C(1|1)}. \tag{29}$$

Then declare

$$\begin{aligned}&\text{target present if } \overline{\Lambda}(t) > \Lambda_T\\ &\text{target not present if } \overline{\Lambda}(t) \le \Lambda_T .\end{aligned}$$

This demonstrates that the integrated target likelihood ratio is a sufficient decision statistic for taking an action to declare a target present or not when the criterion of performance is the minimization of the Bayes' risk.

Target Declaration at a Given Confidence Level. Another approach is to declare a target present whenever its probability exceeds a desired confidence level, p_T. The integrated target likelihood ratio is a sufficient decision statistic for this criterion as well. The prescription is to declare a target present or not according to whether the integrated target likelihood ratio exceeds a threshold, this time given by $\Lambda_T = p_T/(1-p_T)$.

A special case of this is the *ideal receiver,* which is defined as the decision rule that minimizes the average number of classification errors. Specifically, if

$$C(1|1) = 0,\ C(\phi|\phi) = 0,\ C(1|\phi) = 1, \text{ and } C(\phi|1) = 1,$$

then minimizing Bayes' risk is equivalent to minimizing the expected number of miscalls of target present or not present. By (29) this is accomplished by setting $\Lambda_T = 1$, which corresponds to a confidence level of $p_T = 1/2$.

Neyman-Pearson Criterion for Declaration. Another standard approach in the design of target detectors is to declare targets present according to a rule which produces a specified false alarm rate. Naturally, the target detection probability must still be acceptable at that rate of false alarms. In the ideal case, one computes the distribution of the target likelihood ratio with and without the target present and sets the threshold accordingly. Using the Neyman-Pearson approach, we find there is a threshold Λ_T such that calling a target present when the integrated target likelihood ratio is above Λ_T produces the maximum probability of detection subject to the specified constraint on false alarm rate.

3.5 Track Before Detect

The process of likelihood ratio detection and tracking is often referred to as *track before detect*. This terminology recognizes that one is tracking a possible target (through computation of $p(t,\cdot)$) before calling the target present. The advantage of track before detect is that it can integrate sensor responses over time on a moving target to yield a detection in cases where the sensor response at any single time period is too low to call a detection. In likelihood ratio detection and tracking, a threshold is set and a detection is called when the target likelihood ratio surface exceeds that threshold. The state at which the peak of the threshold crossing occurs is usually taken to be the state estimate, and one can convert the target likelihood ratio surface to a probability distribution for the target state.

4. EXAMPLES

In this section we present examples of likelihood ratio detection and tracking

4.1 Simple Simulation

In this section we present an example of likelihood ratio detection and tracking which illustrates the basic features of the methodology. This example shows how to integrate sensor responses on a moving target over time so that even though each response is well below the detection threshold, the target log-likelihood ratio in the cell containing the target (which is changing over time) will build up to cross the detection threshold. This allows us to simultaneously declare a detection and estimate the target's state. This feature is particularly important in low signal-to-noise ratio situations.

Target Motion Model. For this example the target moves in a one-dimensional space of position cells represented by **J**, the integers running from $-\infty$ to $+\infty$. The target can have one of nine base velocities where velocity is measured in cells per time step. The state space is

$$S = \mathbf{J} \times \{-3, -2, \ldots, 4, 5\}$$

so that a state $s \in S$ is an ordered pair (i, v) where i is position and v is base velocity. To S we adjoin the state ϕ, target not present, to obtain $S^+ = S \cup \{\phi\}$.

Let $X(t) = (i(t), v)$ denote the target state at time t. The target motion model is given by

$$i(t+1) = i(t) + v + \delta(t) \tag{30}$$

where

$$\delta(t) = \begin{cases} 1 & \text{with probability } 0.1 \\ 0 & \text{with probability } 0.8 \\ -1 & \text{with probability } 0.1 \end{cases} \tag{31}$$

and $\delta(t)$ is independent of $\delta(t')$ for $t' \neq t$. Intuitively, the target has base velocity v, but there is noise in the velocity process so that the velocity varies randomly between $v-1$ and $v+1$.

For the initial distribution on target state, we assume that the probability the target is not present is 0.75. If the target is present, then its position at time 0 has a uniform distribution on $\{1, \ldots, 100\}$, and its base velocity has a uniform distribution over the nine possible velocities. Specifically, the remaining 0.25 probability is spread uniformly over the 900 position-velocity pairs in $\{1, \ldots, 100\} \times \{-3, -2, \ldots, 4, 5\}$. The result is

$$p(0,\phi) = 0.75$$
$$p(0,(i,v)) = \frac{0.25}{900} = \frac{1}{3600} \text{ for } 1 \le i \le 100, \ \ -3 \le v \le 5$$
$$\lambda(0,(i,v)) = \ln \frac{p(0,(i,v))}{p(0,\phi)} = -\ln(2700) = -7.9$$

Sensor Model and Measurement Likelihood Function. The sensor for this example views only the spatial cells with indices between 1 and 100 inclusive. At each time step, the sensor receives a normally distributed response $Y(i)$ from cell *i* for $1 \le i \le 100$. If the target is in cell *j*, for $1 \le j \le 100$, then $Y(j)$ has mean *u* and standard deviation 1. The target signal spills over to the adjacent cells so that the responses from those cells have mean *u*/2 and standard deviation 1. In all other cells, the response has mean 0 and standard deviation 1. Thus if the target is in cell *j*,

$$Y(i) \sim \begin{cases} \mathcal{N}(0,1) & \text{for } i<j\text{-1 or } i > j+1 \\ \mathcal{N}(\tfrac{1}{2}u,1) & \text{for } i=j\text{-1 or } j\text{+1} \\ \mathcal{N}(u,1) & \text{for } i = j \end{cases}$$

Let $\mathbf{y} = (y_1, \ldots, y_{100})$ be the vector of responses obtained at a given time. Then the likelihood function is

$$L(\mathbf{y} \,|\, (j,v)) = \left(\prod_{\substack{i<j-1 \\ i>j+1}} \eta(y_i, 0, 1) \right) \eta(y_{j-1}, \tfrac{1}{2}u, 1)\, \eta(y_j, u, 1)\, \eta(y_{j+1}, \tfrac{1}{2}u, 1)$$

where $\eta(\cdot, \mu, \sigma^2)$ is the density function for a $\mathcal{N}(\mu, \sigma^2)$ distribution. Note that the above product omits factors with subscripts less than 1 or greater than 100. The likelihood function does not depend on velocity. The measurement log-likelihood ratio is

$$\ln \mathcal{L}(\mathbf{y} \,|\, (j,v)) = \frac{uy_{j-1}}{2} + uy_j + \frac{uy_{j+1}}{2} - \frac{3u^2}{4} \qquad (32)$$

Signal-to-Noise Ratio. To compute the signal-to-noise ratio, we restrict our attention to the three cells containing the signal. In those cells we are dealing with a signal that has a three-dimensional Gaussian distribution with

$$\text{mean } \mu_1 = (u/2, u, u/2)^T \text{ and covariance } \Sigma = \mathbf{I}$$

where **I** is the three-dimensional identity matrix. The noise in those cells has a Gaussian distribution with mean $\mu_0 = (0,0,0)^T$ and covariance, **I**. Thus the signal-to-noise ratio is

$$SNR = \mu_1^T \Sigma^{-1} \mu_1 = \frac{3u^2}{2} \tag{33}$$

4.1.1 Simulated Detection and Tracking Results

This section presents the results of applying likelihood ratio tracking and detection to simulated data. To produce the data, we set the target signal level $u = 2$. This yields an $SNR = 6$ (7.8 dB). The target started at $i = 30$ at time 0 and moved with constant velocity $v = 1$. We used a simulation to produce sensor responses in the 100 cells viewed by the sensor for 16 time periods. Using these responses, we calculated the measurement log-likelihood ratio function in (32) for $t = 0,\ldots,15$.

Since no mass moves from ϕ to S or from S to ϕ, we used the Simplified Likelihood Ratio Recursion with (26) in place of (23) to calculate the target log-likelihood ratio function $\lambda(t,\cdot)$ for $t = 0,\ldots,15$. To perform the step in (26), we exponentiated $\lambda(t_{k-1},s)$ to obtain $\Lambda(t_{k-1},s)$, performed the motion update to get $\Lambda^-(t_k,\cdot)$, and took logarithms to obtain $\lambda^-(t_k,s)$. Finally, we applied (28) to compute $\lambda(t_k,s)$.

Figures 4 to 7 show the target log-likelihood ratio surface $\lambda(t,\cdot)$ at times $t = 0$, 5, 10, and 15, respectively. Note, we show this surface only for the spatial cells with indices between 1 and 100 inclusive. The surface at time 0 (Figure 4) shows the effect of the observation at time 0. This surface is produced by adding the measurement log-likelihood ratio to the prior target log-likelihood ratio. Since the prior target log-likelihood ratio is a constant, equal to –7.9 for all states (except ϕ), the surface at time 0 is simply the measurement log-likelihood ratio for time 0 shifted down by 7.9 log-likelihood units. Since the likelihood function does not depend on velocity, the surface at time 0 shows no velocity dependence. The target's state is (30,1) at time 0. This surface is typical of the measurement log-likelihood ratio functions produced in the simulation. There are many peaks with roughly the same height as the one at the target state.

Figure 5 shows the target log-likelihood ratio surface at $t = 5$. There is a peak developing in the vicinity of the target state at (35,1), but there are many noise peaks that are almost as high. By time $t = 10$, the picture (Figure 6) has clarified dramatically, the largest peak is clearly at the target state (40,1). The height of this peak is 29 log-likelihood units. There is still a substantial amount of background noise in the surface, although the peak levels are much below the one over the target state. At time $t = 15$ (Figure 7), the peak at the target state is equal to 31 log-likelihood units and most of the noise peaks have dropped below –30 log-likelihood units. Because of the "plant noise" in the velocity model (see (31)), the likelihood ratio at the target state is dispersed by each time step of motion. This produces the spread in the log-likelihood ratio peak at the target state that we see in Figures 6 and 7.

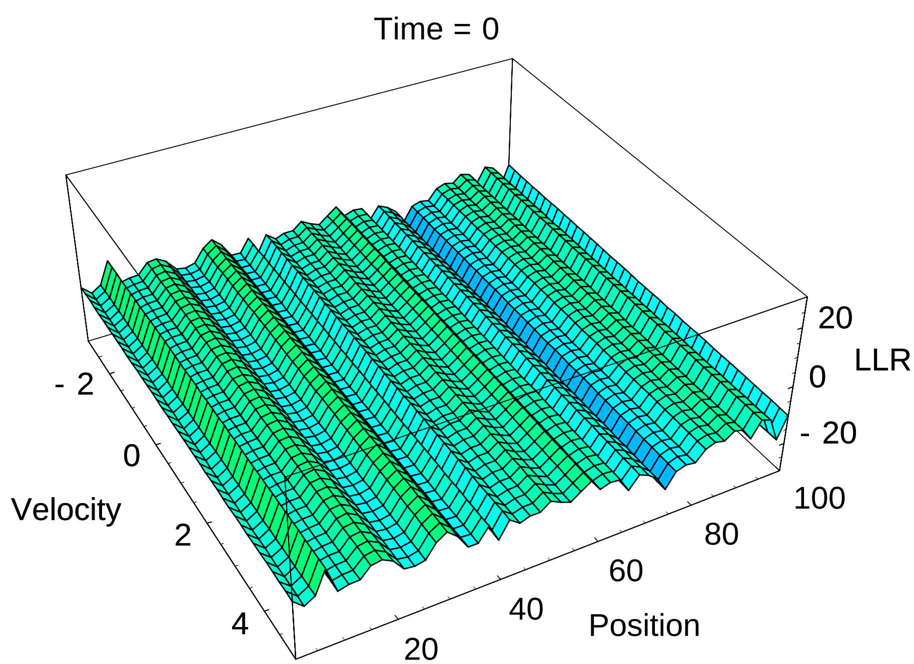

Figure 4 Target Log-Likelihood Ratio Surface at t = 0.

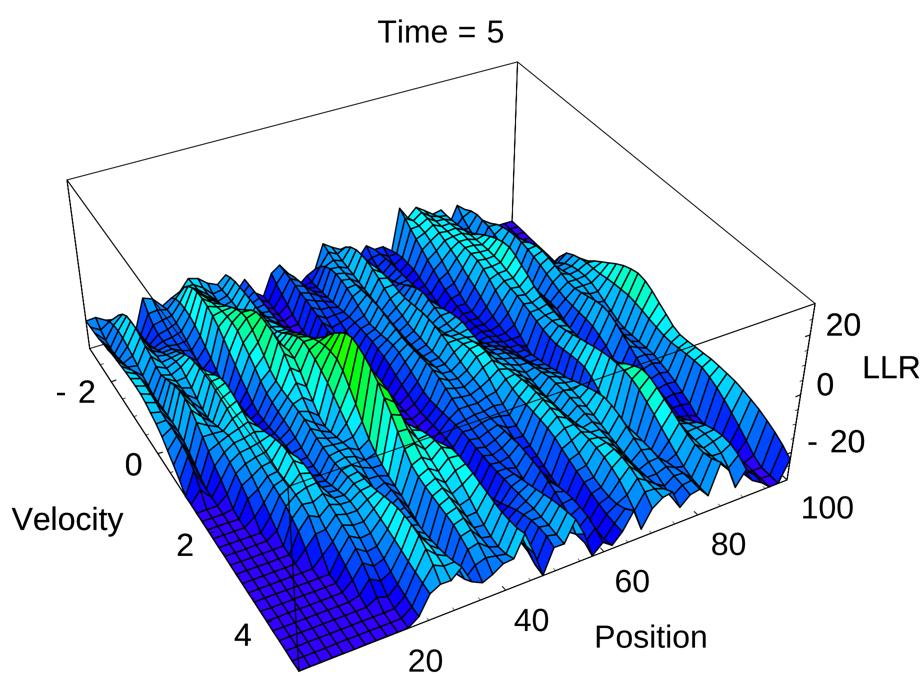

Figure 5 Target Log-Likelihood Ratio Surface at t = 5.

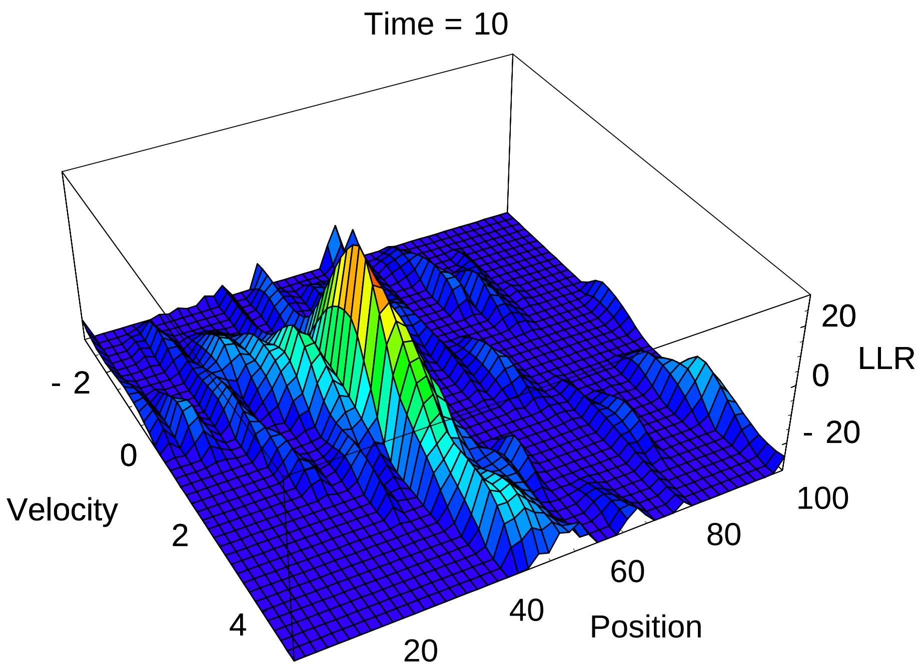

Figure 6 Target Log-Likelihood Ratio Surface at $t = 10$.

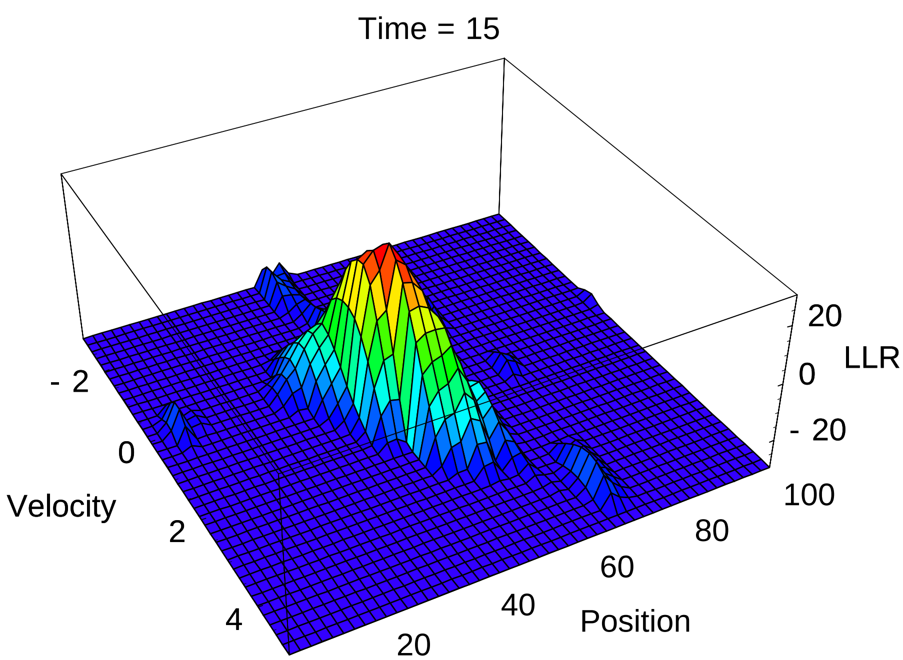

Figure 7 Target Log-Likelihood Ratio Surface at $t = 15$.

4.1.2 Comparison to Matched Filter Detection

To call detections and estimate target state, we set a threshold λ_T and call a detection (target present) whenever the target log-likelihood ratio surface rises above λ_T. The state at the peak of the surface can be used as a point estimate for the target state. In the case of Figure 7, we could provide a probability distribution for the target's state by fitting a bivariate normal distribution to the peak transformed to likelihood ratio units. This is done by fitting a quadratic to the peak in the target log-likelihood ratio surface.

In this section we compare the effectiveness of likelihood ratio detection and tracking to a more naïve detection scheme which is similar to matched filter processing. The matched filter for the signal given the target is located in cell *j* is

$$H(j) = \frac{uy_{j-1}}{2} + uy_j + \frac{uy_{j+1}}{2} \tag{34}$$

As an alternate to computing the target log-likelihood ratio surface at each time, we could compute the matched filter output for position *j*, for $j = 1, \ldots, 100$ from the sensor responses at each time and declare a detection (and target position estimate) whenever the response for some position *j** exceeds a specified threshold λ_M.

Performance of Likelihood Ratio Detection and Tracking. Section 6.2.2 of Stone *et al.* (1999) compares the performance of this matched filter detection methodology, which does not integrate sensor responses over time, to that of likelihood ratio detection and tracking. The analysis in Stone *et al.* showed that for a given threshold setting, the likelihood ratio detection methodology produces a 0.93 probability of detection at a specified false alarm rate. In order to obtain that same detection probability with the matched filter detector, one has to suffer a false alarm rate that is higher by a factor of 10^{18}. This is an example of the impressive increase in performance that can be produced by using likelihood ratio detection and tracking. This increase in performance allows one to detect and track targets at low signal-to-noise ratios.

4.2 Periscope Detection

We now present an example of likelihood ratio detection and tracking that involves a single target and a high rate of false alarms and clutter. The objective is to use a high-resolution shipboard radar to detect submarine periscopes exposed for only a few seconds at a time within a 10 mile radius of a ship.

The radar in question has a 1 ft range resolution, 2 degree beamwidth, and a 5 Hz scan rate. It is assumed that a periscope will be exposed on the order of 10 seconds, although this information is not directly used by the tracking system. Because the radar is mounted on a ship and is looking out to ranges of up to 10 miles, the grazing angle of the radar signal is very low to the surface of the ocean. A high resolution radar with a low grazing angle encounters significant clutter from breaking waves which generate substantially higher returns than the mean ambient level. These high-intensity clutter spikes produce a high clutter rate and potentially a high false alarm rate.

The statistical behavior of this spiky clutter is described by a two-scale model. In this model the received radar signal is represented as a compound process where a fast speckle process is modulated by a slower process describing the scattering features in a radar cell. Over short time intervals (up to approximately 250 ms) the intensity observed in a fixed range cell is Rayleigh distributed. Over longer times, the mean of the Rayleigh component follows a gamma distribution whose shape parameter is a function of the radar and ocean parameters such as grazing angle, resolution cell size, frequency, look direction, and sea state.

Since the radar is looking for a submarine periscope, it is reasonable to assume that there is at most one target at any time within a 10 mile radius of the ship. The result is that we have a single (or no target) problem with a potentially large number of false alarms. To tackle this problem Stone *et al*. (1997) used a Likelihood Ratio Tracker (LRT). The LRT employed consists of a clutter tracker and a target tracker. The clutter tracker estimates the mean clutter level (i.e., the mean of the Rayleigh distribution mentioned above) in each range cell, every 1/5 of a second using the intensity of the radar returns.

Target Tracker. Each beam of the radar is treated separately. The nominal periscope is up for only a short period of time, and the chance of transiting from one beam to another is small, especially with overlapping beams. Within each beam, the target's state space is two-dimensional: *range* and *range rate*. All quantities are measured relative to the radar. The tracker does not try to estimate motion orthogonal to the look direction because with the scales involved it is difficult to use a single radar to "triangulate" the target.

The range space was discretized into 1 ft cells to match the radar range resolution. The range rate component was discretized into 100 cells with a spacing of 0.1 m/sec to cover the range from roughly –10 kn to +10 kn. The initial distribution was chosen with probability 5×10^{-6} of the target being in the state space (i.e., the periscope being up). This probability is uniformly distributed over the target state space. The initial probability of no target present in the state space is $1-5\times10^{-6}$. To allow for the possibility that a periscope exposure commences during a scan, there is probability 5×10^{-8} of a periscope appearing in a given cell per time period (0.2 seconds).

For each scan, the measurement used by this tracker is the set of observed intensities of the radar return from each range cell. The observed intensity and the estimate of the mean clutter level for a cell are used to compute a measurement likelihood ratio statistic in that cell. This statistic is the ratio of the probability of receiving the return given the target is in the cell (i.e., the periscope is exposed) to the probability of receiving the return given no target present. The two probabilities are conditioned on the estimate of the mean clutter level in the cell. This measurement likelihood ratio function is combined by pointwise multiplication with the motion updated likelihood ratio for the target to produce the posterior likelihood ratio function over the target state space. The resulting likelihood ratio surface is then updated for motion (using the motion model described below) to form the motion updated likelihood ratio for the next time increment.

Over the short period of exposure of a periscope, the tracker assumes a constant course and speed for the target. This produces a constant range rate for the target. Let $(r,\dot{r})$ represent a range and range-rate cell in the target state space. The likelihood ratio in cell $(r,\dot{r})$ is displaced to the cell $(r+\Delta t\,\dot{r},\dot{r})$ over a time interval Δt. In this motion model, each velocity hypothesis can be treated independently.

If the radar observations are consistent with a velocity hypothesis, peaks will develop in the likelihood ratio function at position-velocity cells consistent with that hypothesis. Viewed over time, the peaks will lie along a straight line corresponding to the target's track in range versus time. Peaks that occur as noise and do not form according to a velocity hypothesis will not be reinforced. They will tend to average to the background noise level. Typically one sets a threshold level and calls a detection whenever a peak exceeds the threshold. The state at which the peak occurs is used as the estimate of the target state. The process of integrating the likelihood ratio functions over time according to velocity hypotheses provides much of the power of the likelihood ratio tracker methodology.

Example. The likelihood ratio tracker described above was applied to simulated clutter data with an injected target signal to produce the results shown below. Figure 8 shows the radar scan data (intensity on logarithmic scale) for 30 seconds of data and for a 500 ft range interval. In the figure, time increases down the page and range increases from left to right. Each horizontal line in the figure represents a single scan, and there are 150 scans of data shown. The color scale had been adjusted to vary from blue for the minimum value to red for the maximum.

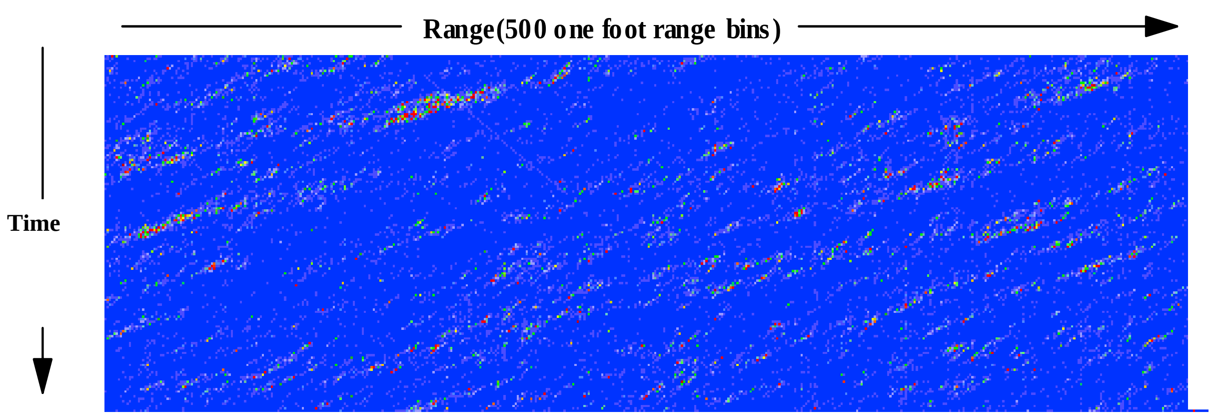

Figure 8 Radar Scan Data for 30 seconds (150 scans)

In Figure 8, a target has been injected amongst the clutter for a 10-second interval. This figure shows the clutter problem faced by the radar in this situation. The clutter patches show up as darker areas moving toward the radar. The patches appear predominantly along the crests of long waves. It is very difficult to see the target track in this view. A single uniform threshold that was set low enough to capture a substantial number of the target returns would also let pass a large number of false alarms. It is clear that any simple thresholding scheme will either call an overwhelming number of false alarms or provide a very low detection probability.

Figure 9 shows the output of the likelihood ratio tracker. The axes and dimensions in this figure are identical to those in Figure 8. Each horizontal line represents the marginal posterior likelihood ratio function calculated through that scan. Each dot on the line shows the marginal likelihood ratio in that range at that time. The marginal likelihood ratio is obtained by adding the likelihood ratio corresponding to all velocity hypotheses in that range cell. The marginal is shown for convenience of display only. The actual comparison of peak values to a threshold is performed on the likelihood ratios as a function of both position and velocity.

The results shown in Figure 9 are striking. The target track, which is not visible in Figure 8, is shown clearly in this figure. It has a peak likelihood ratio value of 4×10^7 compared to the value of 7.6 for the next highest peak in the figure. If we draw the path of a target with a constant velocity (range rate) in the figure, it will appear as a straight line. The line will slant to the right if range is increasing and to the left if it is decreasing. The likelihood peaks line up very well along the target path to form an almost straight line. The peaks disappear when the periscope submerges.

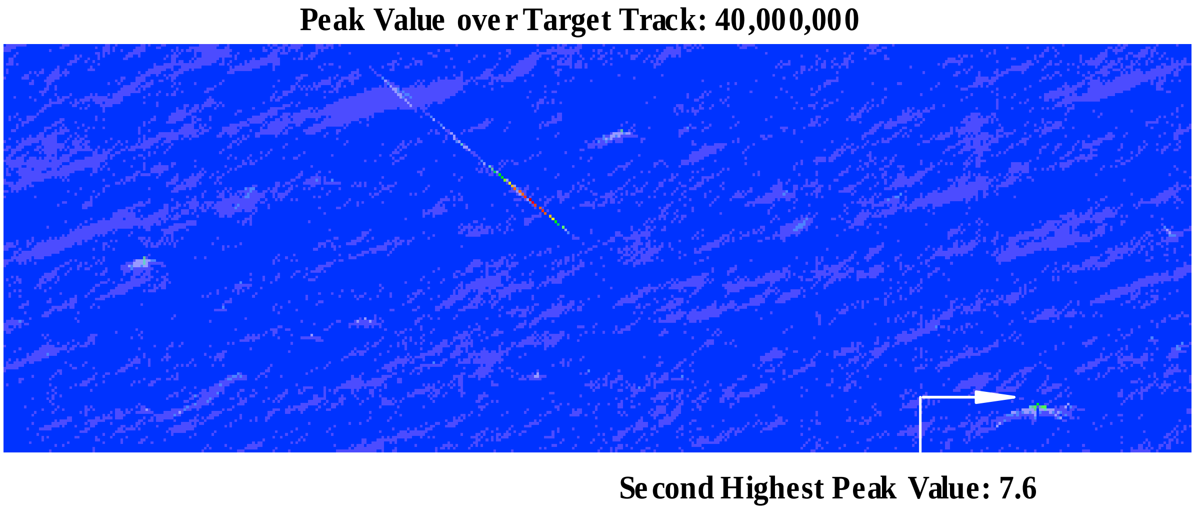

Figure 9 Output of the likelihood ratio tracker.

An expanded view of the likelihood peaks is shown in the left-hand side of Figure 10. Here we can see the buildup of the likelihood ratio along the target track. (Figure 10 employs a gray scale with white indicating the minimum and black the maximum.) The right-hand side of the figure shows the marginal likelihood ratio in velocity (range rate) for the same time period as shown in Figure 9. Here we can see a clear peak at the correct range rate. This is the one corresponding to the line of peaks shown in Figures 9 and 10.

Notice that in this example the tracker uses unthresholded sensor output. Specifically, it does not rely on a thresholding scheme to call contacts and produce sensor measurements. This tracker is based on sensor output rather than contacts. As a result it can accumulate the sensor information from a number of sensor responses over time, all of which may be below a threshold value, until the cumulative result crosses a threshold and allows the tracker to simultaneously call a detection and provide a track estimate.

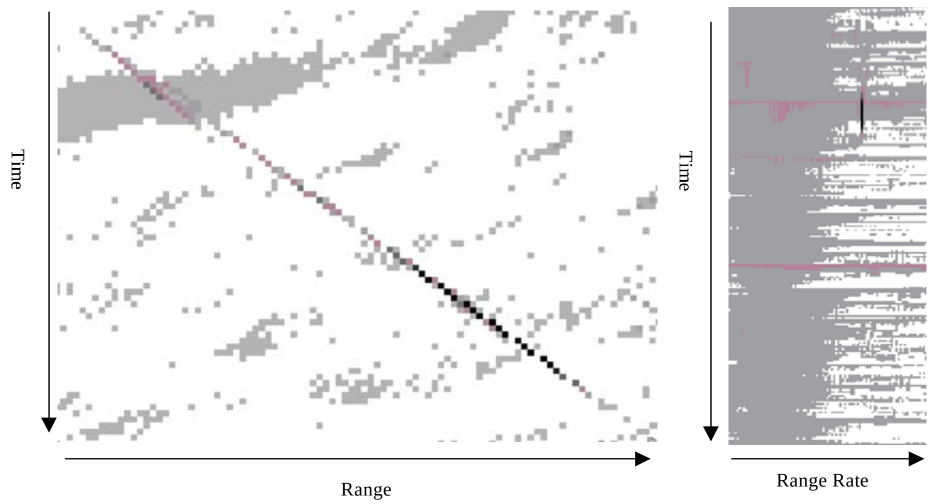

Figure 10 Expanded view of likelihood ratio peaks (left) and velocity likelihood marginal (right).

Since the tracker has been accumulating sensor responses along many possible tracks, it has, in some sense, been performing tracking before detection. In the example shown here, the tracker implicitly considers over 30 million possible tracks to determine if the cumulative likelihood ratio for any one of them exceeds a specified threshold.

4.3 TENET Example

Musick *et al.* (2001) have explored the applicability of nonlinear tracking techniques to the problem of low SNR tracking, particularly in the case were the sensor data are unthresholded outputs from a grid of pixels. An example is pixelized data from point-target image tracking data.

Sensor Model. Musick *et al.* consider the following sensor model. The image contains M pixels, and the sensor measurement at time k consists of the pixel output vector

$$\mathbf{Y}_k = \left(Y_{k1}, \ldots, Y_{kM}\right).$$

The target occupies only one pixel in the image. If there is no target present at pixel m at time k, then the distribution of the output Y_{km} of pixel m is Rayleigh distributed with density

$$f_0(y_{km}) = y_{km} \exp\left(\frac{-y_{km}^2}{2}\right) \text{ for } y_{km} \geq 0. \tag{35}$$

If the target is located at pixel m at time k, then the pixel output is Rayleigh distributed with density

$$f_1(y_{km}) = \frac{y_{km}}{1+\lambda}\exp\left(\frac{-y_{km}^2}{2(1+\lambda)}\right) \text{ for } y_{km} \geq 0 \tag{36}$$

where $\lambda > 0$. Calculating the divergence between f_0 and f_1 one obtains

$$\frac{\lambda^2}{1+\lambda}$$

which Musick *et al.* call the *signal-to-noise* ratio for this sensor. The distribution of the sensor response in one pixel is assumed to be independent of that in any other pixel.

Likelihood Functions. Let $X(k)$ denote the position (pixel) of the target at time *k*. Then the likelihood function for the sensor response $\mathbf{Y}_k = \mathbf{y}_k = (y_{k1},\ldots,y_{kM})$ is

$$L(\mathbf{y}_k \mid m) \equiv \mathbf{Pr}\left\{\mathbf{Y}_k = \mathbf{y}_k \mid X(k) = m\right\} = f_1(y_{km})\prod_{i\neq m} f_0(y_{ki}) \text{ for } m = 1,\ldots,M \ ,$$

and the measurement likelihood ratio function is

$$\mathcal{L}(\mathbf{y}_k \mid m) = \frac{f_1(y_{km})}{f_0(y_{km})} = \frac{1}{1+\lambda}\exp\left(\frac{\lambda y_{km}^2}{2(1+\lambda)}\right) \text{ for } m = 1,\ldots,M \ . \tag{37}$$

Challenge Problem. Musick and Greenewald (2001) have posed a challenge problem to promote the development of effective numerical techniques for nonlinear tracking applications. The problem is called TENET which is an acronym for **TE**chniques for **N**onlinear **E**stimation of **T**racks. As part of this challenge they have developed a simulation (see https://www.tenet.vdl.afrl.af.mil/) which generates target tracks according to the motion model described below and sensor responses according to the pixel distributions described above. Using this simulation, Musick and Greenewald compared the performance of a nonlinear particle-filter tracker and a cellular discrete nonlinear tracker based on the Alternating Direction Implicit (ADI) finite-difference method. They found that the performance of the particle filter was generally superior to the ADI tracker.

The simulation produces sensor output consisting of scenes of 256 by 256 pixels at each time. The trackers are given the approximate location of the target at time 0 within a gate of 10 by 10 pixels. The challenge is to see how well the tracker can localize and maintain lock on the target given this initial "detection."

Motion Model. Target state, $X = (x_1, v_1, x_2, v_2)$, is 4-dimensional where x_1 and x_2 are the position coordinates and v_1 and v_2 the velocity coordinates. The target's motion is a continuous time diffusion sampled at a discrete set of times $t_k = k\Delta$ for some $\Delta > 0$. The target motion is governed by the following equation:

$$X(t_0) = X(0)$$
$$X(t_k) = X^T(t_{k-1})\mathbf{F}(\Delta) + V_k(\Delta) \text{ for } k = 1, 2\ldots$$

where **F**

$$\mathbf{F}(\Delta) = \begin{pmatrix} 1 & \Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and the $V_k(\Delta)$, $k = 1, 2, \ldots$, are independent 4-dimensional Gaussian random variables with mean $(0,0,0,0)$ and covariance

$$\mathbf{E}\left[V_k(\Delta)V_k(\Delta)^T\right] = q\begin{pmatrix} \frac{1}{3}\Delta^3 & \frac{1}{2}\Delta^2 & 0 & 0 \\ \frac{1}{2}\Delta^2 & \Delta & 0 & 0 \\ 0 & 0 & \frac{1}{3}\Delta^3 & \frac{1}{2}\Delta^2 \\ 0 & 0 & \frac{1}{2}\Delta^2 & \Delta \end{pmatrix}.$$

In the example below we take $q = 5 \times 10^{-4}$ which produces target paths with almost constant velocity.

Example. In the challenge problem, it is assumed that the target has been detected and the location of the detected target is used to start the tracking problem. For the example that we present below, we shall back up one step and see whether we can use a Likelihood Ratio Tracker (LRT) to detect the target. The detection called by the LRT could be used to initiate a tracker.

For our example we took the scenes to consist of 61 x 61 pixels with the field of view as shown in Figure 11. There are 10 pixels per unit length. The target's initial position is $(0,0)$ with velocity $(0.2,0)$. (This is unknown to the LRT.) The target's path is generated from the motion model above in discrete time for $t_k = k$ for $k = 0,\ldots,10$. The signal-to-noise ratio is 10 dB ($\lambda = 10.9$). At each time the pixel response in the pixel containing the target is obtained by an independent draw from the density f_1 in (36), and the response in pixels not containing the target are obtained by independent draws from f_0 in (35).

We generated 100,000 sample points at time 0 from a distribution that is uniform over $[-1.5, 4.5] \times [-3.0, 3.0]$ for position and uniform over $[-0.5, 0.7] \times [-0.6, 0.6]$ for velocity. We took the prior probability of a target being present in the field of regard to be 0.1 so that the initial likelihood ratio on a each point is

$$\Lambda(0) = \frac{0.1 \times 10^{-5}}{0.9} = 1.11 \times 10^{-6} \; (\ln \Lambda(0) = -13.7).$$

For the LRT, we computed the measurement likelihood ratio for each sample point from (37) using the response y_{km} in the pixel closest to the sample point. The cumulative likelihood ratio for this point was then multiplied by the measurement likelihood ratio. After this, the points were resampled to split the ones with high likelihood ratio and delete ones with low values. Each of the sample points was then motion updated according to the motion model described above. The results are shown in Figures 11 – 20 below.

For times $t = 0, 2, 4, 6,$ and 8 the figures show the posterior position log likelihood ratio surface and the posterior velocity log likelihood ratio surface. The white star in the figures marks the target's true position or velocity. At time 0 (Figures 11 and 12.) the first sensor measurement has been processed, and we can see that there is very little information about the target's location or velocity. The same is true for time 2 (Figures 13 and 14.) At time 4 (Figures 15 and 16.) we can see the log likelihood ratio surface beginning to peak up at the target's position and velocity. By times 6 and 8 (Figures 17 – 20), we can see a strong peak in the log likelihood ratio surface near the target's position and velocity.

5. REFERENCES

Blackman, S. S. and Popoli, R. (1999) *Design and Analysis of Modern Tracking Systems,* Artech House, Boston.

Doucet, A., de Freitas, N., Gordon, N (2001) *Sequential Monte Carlo Methods in Practice*. Springer – Verlag, New York.

Jazwinski, A. H. (1970) *Stochastic Processes and Filtering Theory,* Academic Press, New York.

Musick, S., Greenewald, J., Kreuchner, C, and Kastella, K. (2001) "Comparison of Particle Method and Finite Difference Nonlinear Filters for Low SNR Target Tracking" *Conference Proceedings of Fusion 2001,* August 6 –10, 2001.

Musick, S. H., and Greenewald, J. (2001) *TENET Simulator User's Guide,* Air Force Research Laboratory, AFRL/SNAT, Wright-Patterson AFB, OH. July 2001.

Stone, L. D., Barlow, C. A., and Corwin, T.L. (1999) *Bayesian Multiple Target Tracking*, Artech House, Boston.

Stone, L. D., M. V. Finn, and C. A. Barlow, (1997) "Uncluttering the Tactical Picture," *Proceedings of the 1997 IRIS National Symposium on Sensor and Data Fusion*, Fort Belvoir, VA: Infrared Information Analysis Center of Defense Technical Information Center, Vol. II, pp. 127–146.

Van Trees, H. L. (1967) *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*, John Wiley, New York.

Figure 11. Position Log Likelihood Ratio at $t = 0$

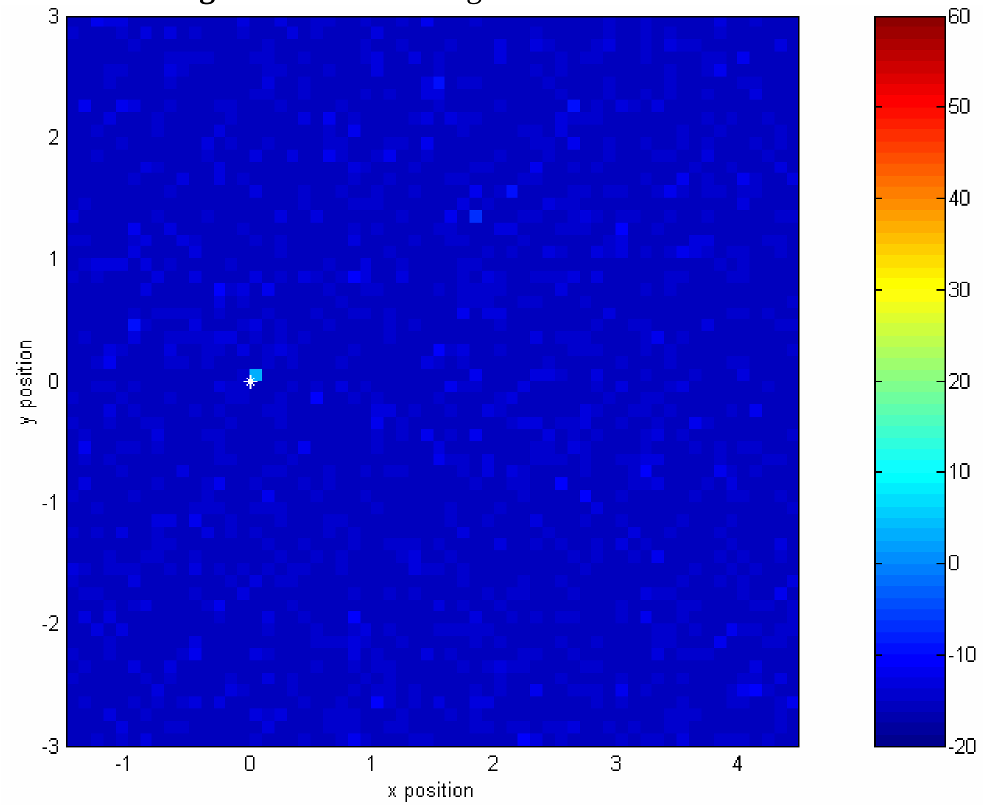

Figure 12. Velocity Log Likelihood Ratio at $t = 0$

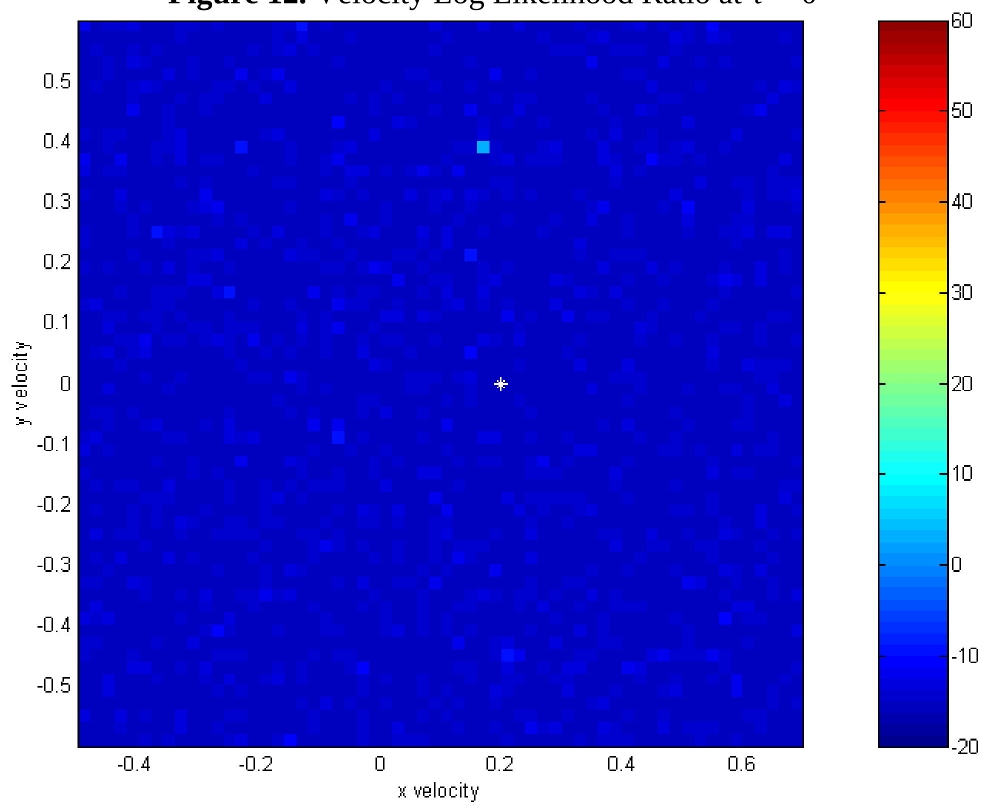

Figure 13. Position Log Likelihood Ratio at $t = 2$

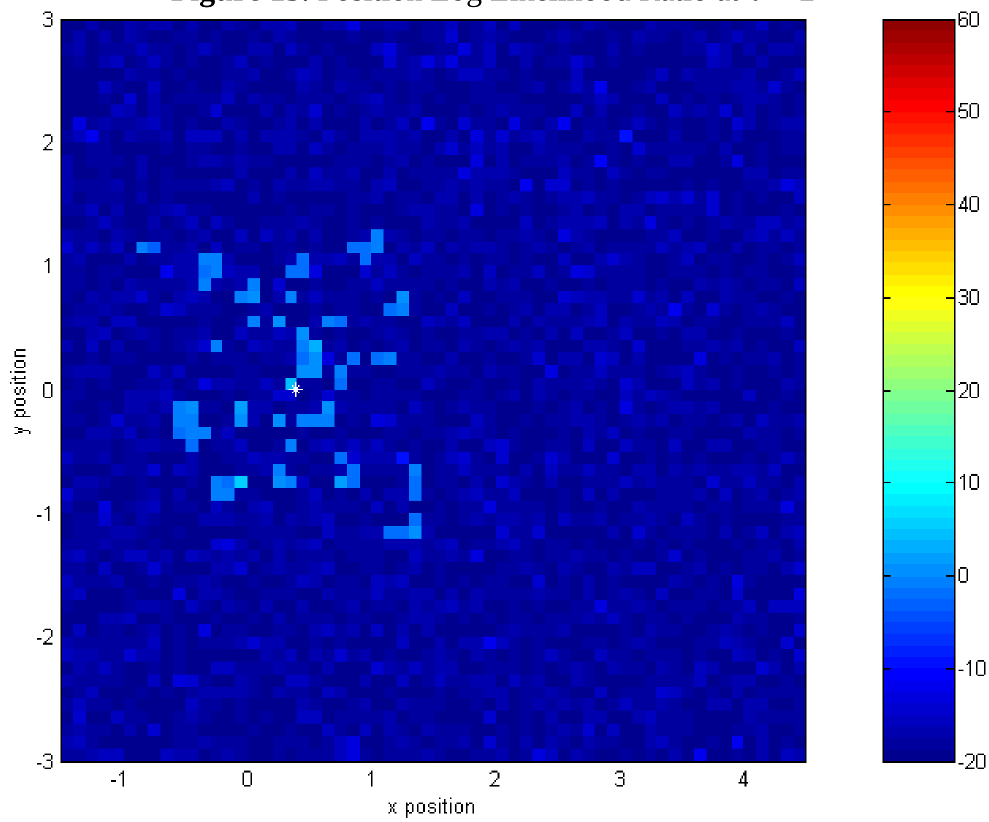

Figure 14. Velocity Log Likelihood Ratio at $t = 2$

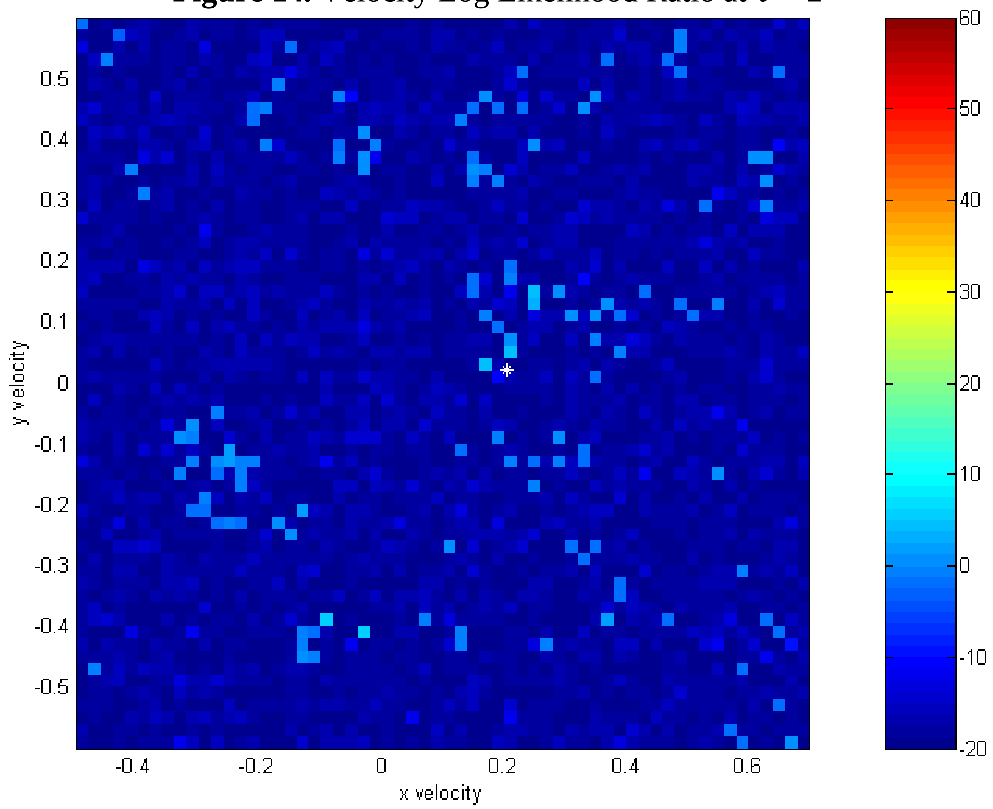

Figure 15. Position Log Likelihood Ratio at $t = 4$

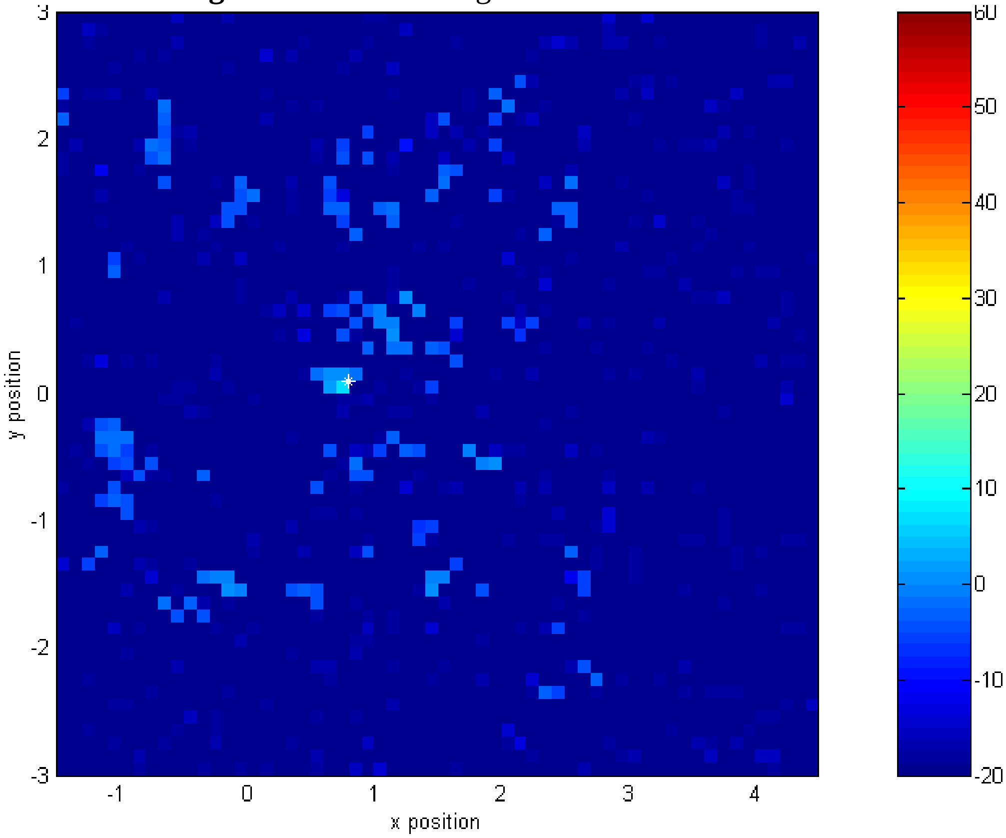

Figure 16. Velocity Log Likelihood Ratio at $t = 4$

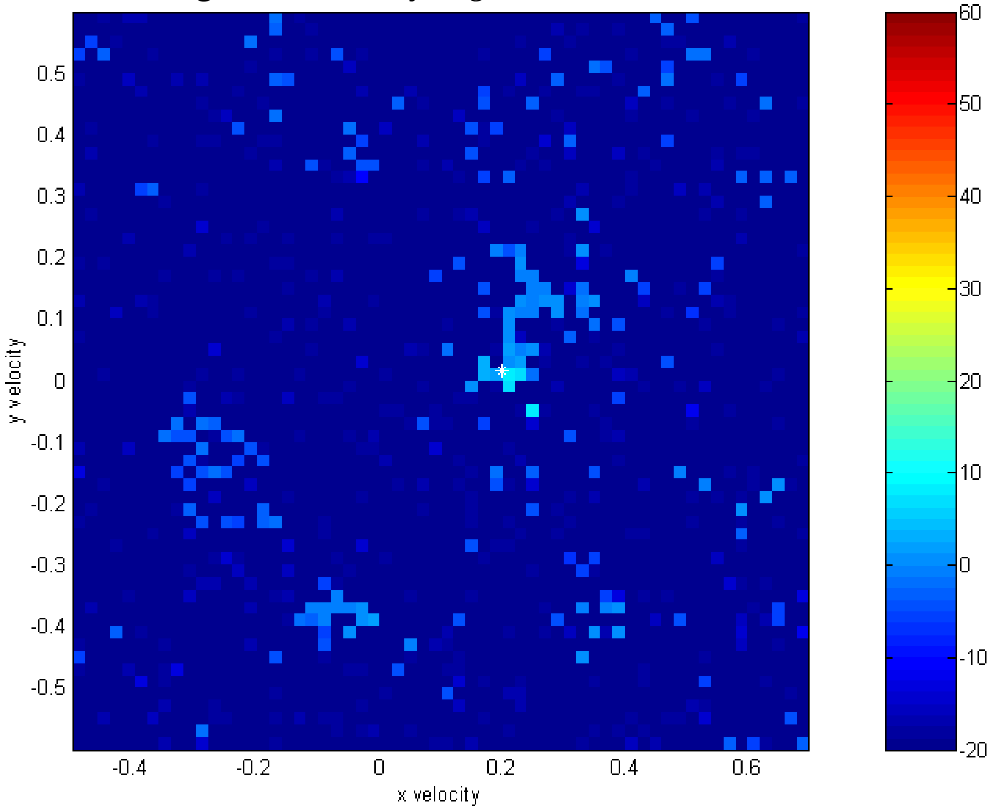

Figure 17. Position Log Likelihood Ratio at $t = 6$

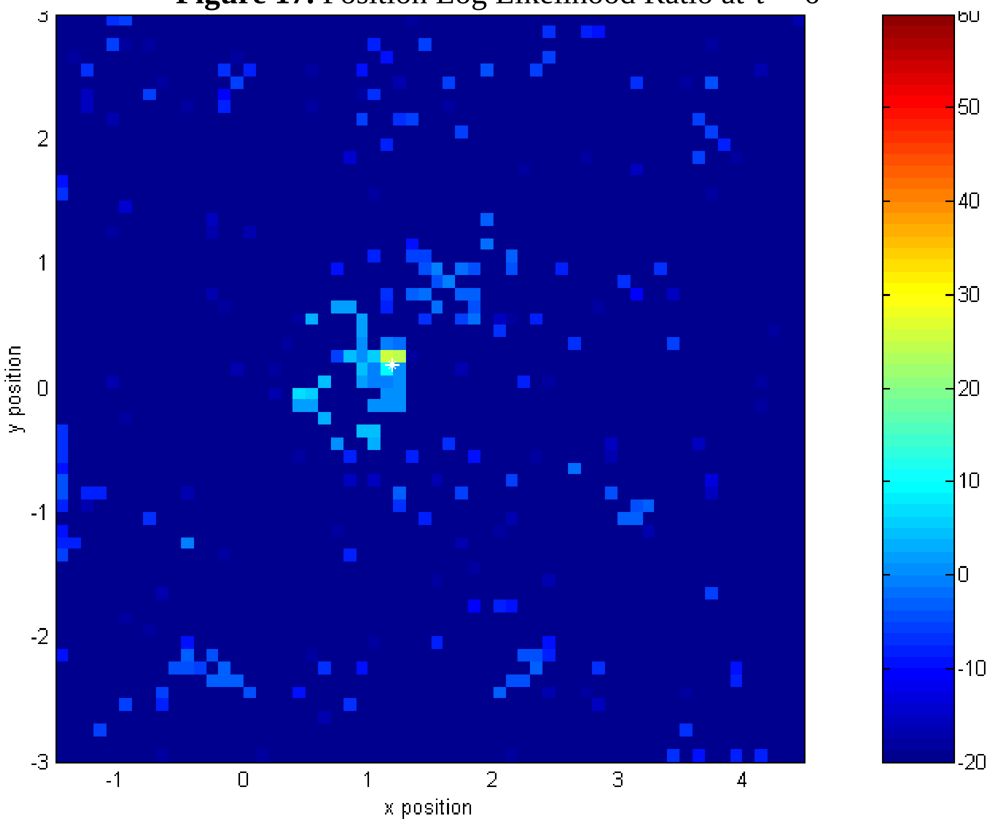

Figure 18. Velocity Log Likelihood Ratio at $t = 6$

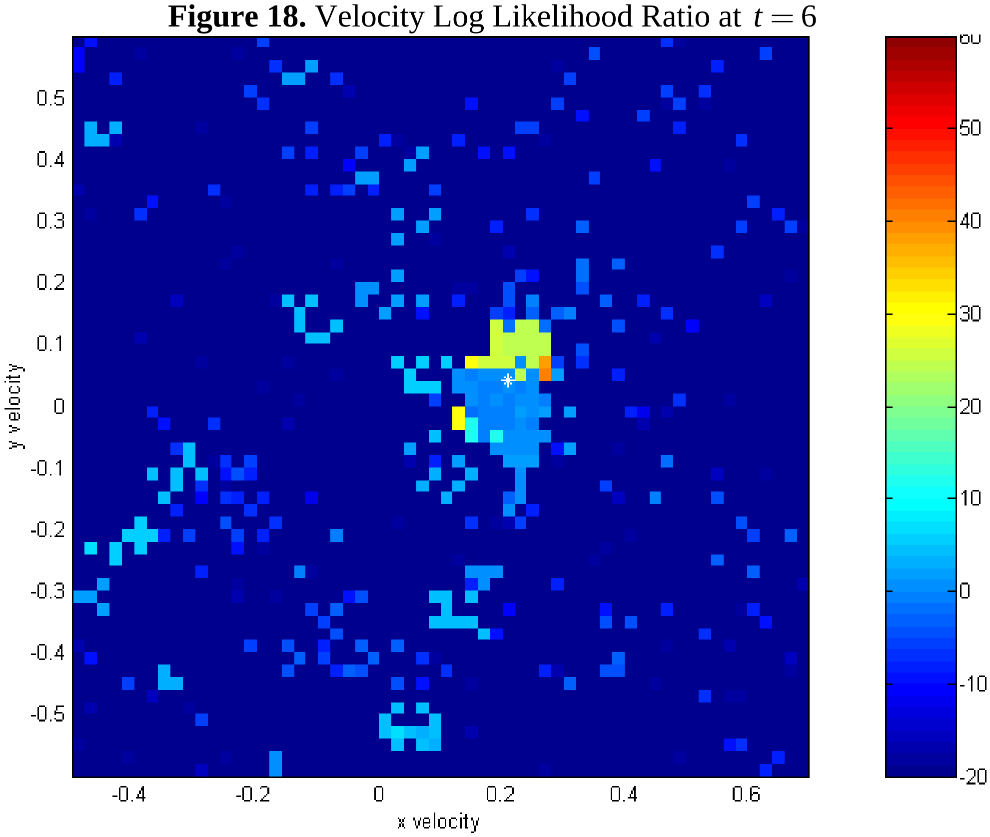

Figure 19. Position Log Likelihood Ratio at $t = 8$

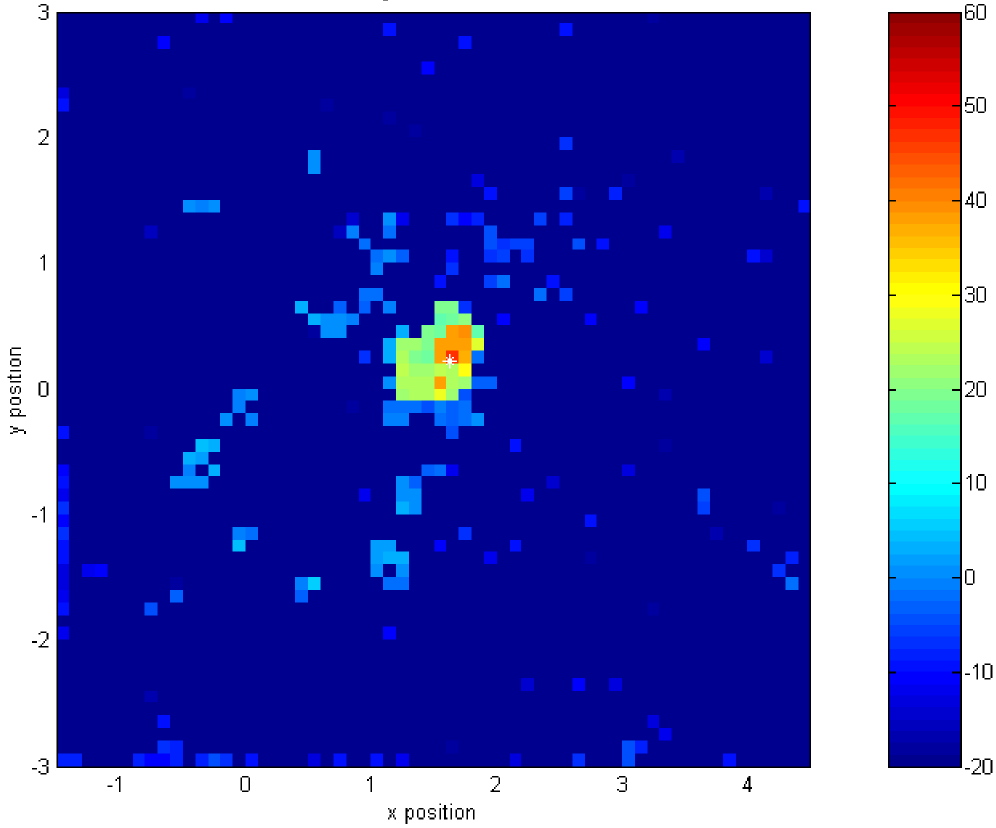

Figure 20. Velocity Log Likelihood Ratio at $t = 8$